

MLLNVLRICI	IVCLVNDGAG	KHSEGRERTK	TYSLNSRGYF	40
RKERGARRSK	ILLVNTKGLD	EPHIGHGDFG	LVAELFDSTR	80
THTNRKEPDM	NKVKLFSTVA	HGNKSARRKA	<u>YNGSRRNIFS</u>	120
RRSFDKRNT	VTEKPGAKMF	WNNFLVKMNG	APONT <u>SHGSK</u>	160
AQEIMKEACK	TLPFTQNIVH	ENCDRMVION	NLCFGKCISL	200
HVPNQODRRN	TCSHCLPSKF	TLNHLTL <u>NCT</u>	GSKNVVKVVM	240
MVEECTCEAH	KSNFHQTAQF	NMDTSTTLHH		270

Figure 1. Deduced amino acid sequence of *Xenopus cerberus* protein. SEQ ID NO:1.

20250323 14:44

Figure 2. Nucleotide sequence of the full-length cerberus DNA derived from the *Xenopus* organizer. The sense strand is on top (in the 5' to 3' direction) and the antisense strand on the bottom line (on the opposite direction). SEQ ID NO:2.

GAATTC	CCAG	CAAGTCGCTC	AGAAACACTG	CAGGGTCTAG	ATATCATACA	ATGTTACTAA	60
CTTAAG	GGGTC	GTTCAGCGAG	TCTTTGTGAC	GTCCAGATC	TATAGTATGT	TACAATGATT	
ATGTACTCAG	GATCTGTATT	ATCGTCTGCC	TTGTGAATGA	TGGAGCAGGA	AAACACTCAG		120
TACATGAGTC	CTAGACATAA	TAGCAGACGG	AACACTTACT	ACCTCGTCCT	TTTGTGAGTC		
AAGGACGAGA	AAGGACAAAA	ACATATTCAC	TTAACAGCAG	AGGTTACTTC	AGAAAAGAAA		180
TTCTGTCTCT	TTCTGTTTTT	TGTATAAGTG	AATTGTCGTC	TCCAATGAAG	TCTTTTCTTT		
GAGGAGCAGG	TAGGAGCAAG	ATTCTGCTGG	TGAATACTAA	AGGTCTTGAT	GAACCCACAC		240
CTCCTCGTGC	ATCCTCGTTC	TAAGACGACC	ACTTATGATT	TCCAGAACTA	CTTGGGGTGT		
TTGGGCATGG	TGATTTTCGC	TTAGTAGCTG	AACTATTTGA	TTCCACCAGA	ACACATACAA		300
AACCCGTACC	ACTAAAAGCG	AATCATCGAC	TTGATAAACT	AAGGTGGTCT	TGTGTATGTT		
ACAGAAAAGA	GCCAGACATG	AACAAAGTCA	AGCTTTTCTC	AACAGTTGCC	CATGGAAACA		360
TGTCTTTTCT	CGGTCTGTAC	TTGTTTCAGT	TCGAAAAGAG	TTGTCAACGG	GTACCTTTGT		
AAAGTGCAAG	AAGAAAAGCT	TACAATGGTT	CTAGAAGGAA	TATTTTTCCT	CGCCGTTCCT		420
TTTCACGTTT	TTCTTTTCGA	ATGTTACCAA	GATCTTCCTT	ATAAAAAGGA	GCGGCAAGAA		
TTGATAAAAG	AAATACAGAG	GTTACTGAAA	AGCCTGGTGC	CAAGATGTTT	TGGAACAATT		480
AACTATTTTC	TTTATGTCTC	CAATGACTTT	TCGGACCACG	GTTCTACAAG	ACCTTGTTAA		
TTTTGGTTAA	AATGAATGGA	GCCCCACAGA	ATACAAGCCA	TGGCAGTAAA	GCACAGGAAA		540
AAAACCAATT	TTACTTACCT	CGGGGTGTCT	TATGTTGCGT	ACCGTCATTT	CGTGTCTCTT		
TAATGAAAGA	AGCTTGCAAA	ACCTTGTTTT	TCACTCAGAA	TATTGTACAT	GAAAACCTGTG		600
ATTACTTTCT	TCGAACGTTT	TGGAACAAAA	AGTGAGTCTT	ATAACATGTA	CTTTTGACAC		
ACAGGATGGT	GATACAGAAC	AATCTGTGCT	TTGGTAAATG	CATCTCTCTC	CATGTTCCAA		660
TGTCCTACCA	CTATGTCTTG	TTAGACACGA	AACCATTTC	GTAGAGAGAG	GTACAAGGTT		
ATCAGCAAGA	TCGACGAAAT	ACTTGTTCCC	ATTGCTTGCC	GTCCAAATTT	ACCCTGAACC		720
TAGTCGTTCT	AGCTGCTTTA	TGAACAAGGG	TAACGAACGG	CAGGTTTAAA	TGGGACTTGG		
ACCTGACGCT	GAATTGTA	GGATCTAAGA	ATGTAGTAAA	GGTTGTCATG	ATGGTAGAGG		780
TGGACTGCGA	CTTAACATGA	CCTAGATTCT	TACATCATTT	CCAACAGTAC	TACCATCTCC		
AATGCACGTG	TGAAGCTCAT	AAGAGCAACT	TCCACCAAAC	TGCACAGTTT	AACATGGATA		840
TTACGTGCAC	ACTTCGAGTA	TTCTCGTTGA	AGGTGGTTTG	ACGTGTCAAA	TTGTACCTAT		
CATCTACTAC	CCTGCACCAT	TAAAGGACTG	CCATACAGTA	TGGAAATGCC	CTTTTGTGTTG		900
GTAGATGATG	GGACGTGGTA	ATTCCTGAC	GGTATGTCAT	ACCTTTACGG	GAAAACAACC		
AATATTTGTT	ACATACTATG	CATCTAAAGC	ATTATGTTGC	CTTCTATTTC	ATATAACCAC		960
TTATAAACAA	TGTATGATAC	GTAGATTTTCG	TAATACAACG	GAAGATAAAG	TATATTGGTG		
ATGGAATAAG	GATTGTATGA	ATTATAATTA	ACAAATGGCA	TTTTGTGTAA	CATGCAAGAT		1020
TACCTTATTC	CTAACATACT	TAATATTAAT	TGTTTACCGT	AAAACACATT	GTACGTTCTA		

MSRTRKVD	SL	LLLAIPGLAL	LLLPNAYCAS	CEPVRIPMCK	SMPWNMTKMP	NHLHHSTQAN	60
AILAIEQFEG	LLTTECSQDL	LFFLCAMYAP	ICTIDFQHEP	IKPCKSVCKER	ARAGCEPILI		120
KYRHTWPESL	ACEELPVYDR	GVCISPEAIV	TVEQGTDSMP	DFSMDSNNGN	CGSGREHCKC		180
KPMKATQKTY	LKNYNYVIR	AKVKEVKVKC	HDATAIVEVK	EILKSSLVNI	PKDTVTLTYN		240
SGCLCPQLVA	NEEYIIMGYE	DKERTRLLLV	EGSLAEKWRD	RLAKVKRWD	QKLRRPRKSK		300
DPVAPIPNKN	SNSRQARS						

Figure 3. Deduced amino acid sequence of Xenopus frazzled protein. SEQ ID NO:3.

[illegible]

GAATTCCTT	TCACACAGGA	CTCCTGGCAG	AGGTGAATGG	TTAGCCCTAT	GGATTTGGTT	60
CTTAAGGGAA	AGTGTGTCT	GAGGACCGTC	TCCACTTACC	AATCGGGATA	CCTAAACCAA	
TGTTGATTTT	GACACATGAT	TGATTGCTTT	CAGATAGGAT	TGAAGGACTT	GGATTTTTAT	120
ACAACATAAA	CTGTGTACTA	ACTAACGAAA	GTCTATCCTA	ACTTCTCTGAA	CCTAAAAATA	
CTAATTCCTGC	ACTTTTAAAT	TATCTGAGTA	ATTGTTTCATT	TTGTATTGGA	TGGGACTAAA	180
GATTAAGACG	TGAAAATTTA	ATAGACTCAT	TAACAAGTAA	AACATAACCT	ACCCTGATTT	
GATAAACTTA	ACTCCTTGCT	TTTGACTTGC	CCATAAACTA	TAAGGTGGGG	TGAGTTGTAG	240
CTATTTGAAT	TGAGGAACGA	AAACTGAACG	GGTATTTGAT	ATTCCACCCC	ACTCAACATC	
TTGCTTTTAC	ATGTGCCCAG	ATTTTCCCTG	TATTCCCTGT	ATTCCCTCTA	AAGTAAGCCT	300
AACGAAAATG	TACACGGGTC	TAAAAGGGAC	ATAAGGGACA	TAAGGGAGAT	TTCATTCGGA	
ACACATACAG	GTTGGGCAGA	ATAACAATGT	CTCGAACAAG	GAAAGTGGAC	TCATTACTGC	360
TGTGTATGTC	CAACCCGTCT	TATTGTTACA	GAGCTTGTTT	CTTTCACCTG	AGTAATGACG	
TACTGGCCAT	ACCTGGACTG	GCGCTTCTCT	TATTACCCAA	TGCTTACTGT	GCTTCGTGTG	420
ATGACCGGTA	TGGACCTGAC	CGCGAAGAGA	ATAATGGGTT	ACGAATGACA	CGAAGCACAC	
AGCCTGTGCG	GATCCCCATG	TGCAAACTTA	TGCCATGGAA	CATGACCAAG	ATGCCCAACC	480
TCGGACACGC	CTAGGGGTAC	ACGTTTAGAT	ACGGTACCTT	GTA CTGGTTC	TACGGGTTGG	
ATCTCCACCA	CAGCACTCAA	GCCAAATGCCA	TCTTGGCAAT	TGAACAGTTT	GAAGGTTTGC	540
TAGAGGTGGT	GTCGTGAGTT	CGGTTACGGT	AGGACCGTTA	ACTTGTCAAA	CTTCCAAACG	
TGACCACTGA	ATGTAGCCAG	GACCTTTTGT	TCTTTCTGTG	TGCCATGTAT	GCCCCCATTT	600
ACTGGTGACT	TACATCGGTC	CTGGAAAACA	AGAAAGACAC	ACGGTACATA	CGGGGGTAAA	
GTACCATCGA	TTTCCAGCAT	GAACCAATTA	AGCCTTGCAA	GTCCGTGTGC	GAAAGGGCCA	660
CATGGTAGCT	AAAGGTGCTA	CTTGGTTAAT	TCGGAACGTT	CAGGCACACG	CTTCCCGGT	
GGGCCGGCTG	TGAGCCCATT	CTCATAAAGT	ACCGGCACAC	TTGGCCAGAG	AGCCTGGCAT	720
CCCGGCCGAC	ACTCGGGTAA	GAGTATTTCA	TGGCCGTGTG	AACCGGTCTC	TCGGACCGTA	
GTGAAGAGCT	GCCCCGTATAT	GACAGAGGAG	TCTGCATCTC	CCCAGAGGCT	ATCGTCACAG	780
CACTTCTCGA	CGGGCATATA	CTGTCTCCTC	AGACGTAGAG	GGGTCTCCGA	TAGCAGTGTC	
TGGAACAAGG	AACAGATTCA	ATGCCAGACT	TCTCCATGGA	TTCAACAAT	GGAAATTGCG	840
ACCTTGTTCC	TTGTCTAAGT	TACGGTCTGA	AGAGGTACCT	AAGTTTGTTA	CCTTTAACGC	
GAAGCGGCAG	GGAGCACTGT	AAATGCAAGC	CCATGAAGGC	AACCCAAAAG	ACGTATCTCA	900
CTTCGCCGTC	CCTCGTGACA	TTTACGTTTC	GGTACTTCCG	TTGGGTTTTT	TGCATAGAGT	
AGAATAATTA	CAATTATGTA	ATCAGAGCAA	AAGTGAAAGA	GGTGAAAGTG	AAATGCCACG	960
TCTTATTAAT	GTTAATACAT	TAGTCTCGTT	TTCACTTTCT	CCACTTTCAC	TTTACGGTGC	
ACGCAACAGC	AATTGTGGAA	GTAAGGAGA	TTCTCAAGTC	TTCCTAGTG	AACATTCTTA	1020
TGCGTTGTG	TTAACACCTT	CATTTCTCT	AAGAGTTTCA	AAGGGATCAC	TTGTAAGGAT	

MLLLFRAIPM LLLGLMVLQT DCEIAQYYID EEEPPGTVIA VLSQHSIFNT TDIPATNFRL	60
MKQFNNSLIG VRES DGQLSI MERIDREQIC RQSLHCNLAL DVVSFSKGHF KLLNVKVEVR	120
DINDHSPHFP SEIMHVEVSE SSSVGTRIPL EIAIDEDVGS NSIQNFQISN NSHFSIDVLT	180
RADGVKYADL VLMRELDREI QPTYIMELLA MDGGVPSLSG TAVVNIRVLD FNDNSPVFER	240
STIAVDLVED APLGYLLEL HATDDDEGVN GEIVYGSTL ASQEVRLFK INSRTGSVTL	300
EGQVDFETKQ TYEFEVQAQD LGPNPLTATC KVTVHILDVN DNTPAITITP LTTVNAGVAY	360
IPETATKENF IALISTTDRA SGSNGQVRCT LYGHEHFKLQ QAYEDSYMIV TTSTLDRENI	420
AAYSLTVVAE DLGFPSLGTK KYITVKVSE NDNAPVFSKP QYEASILENN APGSYITTVI	480
ARDSDSQNG KVNRYLVDK VMGQSLTTFV SLDADSGVLR AVRS LDYEKL KQLDFEIEAA	540
DNGIPQLSTR VQLNLRIVDQ NDNCPVITNP LLNNGSGEVL LPISAPQNYL VFQLKAEDSD	600
EGHNSQLFYT ILRDPSRLFA INKESGEVFL KKQLNSDHSE DLSIVVAVYD LGRPSLSTNA	660
TVKFILTDSE PSNVEVILQ PSAEEQHQID MSIIFIAVLA GGCALLLLAI FVACTCKKK	720
AGEFKQVPEQ HGTCNEERLL STPSPQSVSS SLSQSESCQL SINTESENCV VSSNQEQHQQ	780
TGIKHSISVP SYHTSGWHLN NCAMSGHGS HMGHISTKVQ WAKEIVTSMT VTLILVENQK	840
RRALSSQCRH KPVLTQMNQ QGSDMPITIS ATESTRVQKM GTAHCMKRA IDCLTL	

Figure 5. Deduced amino acid sequence of the *Xenopus* PAPC (paraxial protocadherin) protein. It encodes a member of the cadherin family of transmembrane proteins that has dorsalizing activity when constructs are injected into *Xenopus* embryos. SEQ ID NO:5.

Abstract

GAATTTCCCG	AGATGAACCTC	CTTGAGATTG	TTTTAAATGA	CTGCAGGTCT	GGAAGGATTC	60
CTTAAGGGTC	TCTACTTGAG	GAACCTCTAAC	AAAATTTACT	GACGTCAGAG	CCTTCCCTAAG	
ACATTGCCAC	ACTGTTTCTA	GGCATGAAAA	AACTGCAAGT	TTCAACTTTG	TTTTTGGTGC	120
TGTAACGGTG	TGACAAAGAT	CCGTACTTTT	TTGACGTTCA	AAGTTGAAAC	AAAAACCACG	
AACTTTGATT	CTTCAAGATG	CTGCTTCTCT	TCAGAGCCAT	TCCAATGCTG	CTGTTGGGAC	180
TTGAAACTAA	GAAGTTCTAC	GACGAAGAGA	AGTCTCGGTA	AGGTTACGAC	GACAACCCTG	
TGATGGTTTT	ACAAACAGAC	TGTGAAATTG	CCCAGTACTA	CATAGATGAA	GAAGAACCCC	240
ACTACCAAAA	TGTTTGTCTG	ACACTTTAAC	GGGTCATGAT	GTATCTACTT	CTTCTTGGGG	
CTGGCACTGT	AATTGCAGTG	TTGTCAACAAC	ACTCCATATT	TAACACTACA	GATATACCTG	300
GACCGTGACA	TTAACGTCAC	AACAGTGTTG	TGAGGTATAA	ATTGTGATGT	CTATATGGAC	
CAACCAATTT	CCGTCTAATG	AAGCAATTTA	ATAATTCCCT	TATCGGAGTC	CGTGAGAGTG	360
GTTGGTTAAA	GGCAGATTAC	TTCGTTAAAT	TATTAAGGGA	ATAGCCTCAG	GCACTCTCAC	
ATGGGCAGCT	GAGCATCATG	GAGAGGATTG	ACCGGGAGCA	AATCTGCAGG	CAGTCCCTTC	420
TACCCGTCGA	CTCGTAGTAC	CTCTCCTAAC	TGGCCCTCGT	TTAGACGTCC	GTGAGGGAAG	
ACTGCAACCT	GGCTTTGGAT	GTGGTCAGCT	TTTCCAAAGG	ACACTTCAAG	CTTCTGAACG	480
TGACGTTGGA	CCGAAACCTA	CACCAGTCGA	AAAGGTTTCC	TGTGAAGTTC	GAAGACTTGC	
TGAAAGTGGA	GGTGAGAGAC	ATTAATGACC	ATAGCCCTCA	CTTTCCAGT	GAAATAATGC	540
ACTTTCACCT	CCACTCTCTG	TAATTACTGG	TATCGGGAGT	GAAAGGGTCA	CTTTATTACG	
ATGTGGAGGT	GTCTGAAAGT	TCCTCTGTGG	GCACCAGGAT	TCCTTTAGAA	ATTGCAATAG	600
TACACCTCCA	CAGACTTTCA	AGGAGACACC	CGTGGTCCTA	AGGAAATCTT	TAACGTTATC	
ATGAAGATGT	TGGGTCCAAC	TCCATCCAGA	ACTTTCAGAT	CTCAAATAAT	AGCCACTTCA	660
TACTTCTACA	ACCCAGGTTG	AGGTAGGTCT	TGAAAGTCTA	GAGTTTATTA	TCGGTGAAGT	
GCATTGATGT	GCTAACCAGA	GCAGATGGGG	TGAAATATGC	AGATTTAGTC	TTAATGAGAG	720
CGTAACTACA	CGATTGGTCT	CGTCTACCCC	ACTTTATACG	TCTAAATCAG	AATTACTCTC	
AACTGGACAG	GGAAATCCAG	CCAACATACA	TAATGGAGCT	ACTAGCAATG	GATGGGGGTG	780
TTGACCTGTC	CCTTTAGGTC	GGTTGTATGT	ATTACCTCGA	TGATCGTTAC	CTACCCCCAC	
TACCATCACT	ATCTGGTACT	GCAGTGGTTA	ACATCCGAGT	CCTGGACTTT	AATGATAACA	840
ATGGTAGTGA	TAGACCATGA	CGTCACCAAT	TGTAGGCTCA	GGACCTGAAA	TTACTATTGT	
GCCCAGTGTT	TGAGAGAAGC	ACCATTGCTG	TGGACCTAGT	AGAGGATGCT	CCTCTGGGAT	900
CGGGTCACAA	ACTCTCTTCG	TGGTAACGAC	ACCTGGATCA	TCTCCTACGA	GGAGACCCTA	
ACCTTTTGTT	GGAGTTACAT	GCTACTGACG	ATGATGAAGG	AGTGAATGGA	GAAATTGTTT	960
TGGAAAACAA	CCTCAATGTA	CGATGACTGC	TACTACTTCC	TCACTTACCT	CTTTAACAAA	
ATGGATTTCAG	CACTTTGGCA	TCTCAAGAGG	TACGTCAGCT	ATTTAAAAAT	AACTCCAGAA	1020
TACCTAAGTC	GTGAAACCGT	AGAGTTCTCC	ATGCAGTCGA	TAAATTTTAA	TTGAGGTCTT	

CTGGCAGTGT GACCGTCACA	TACTCTTGAA ATGAGAACTT	GGCCAAAGTTG CCGGTTCAAC	ATTTTGAGAC TAAAACTCTG	CAAGCAGACT GTTCTGTCTGA	TACGAATTTG ATGCTTAAAC	1080
AGGTACAAGC TCCATGTTTCG	CCAAGATTTG GGTTCATAAAC	GGCCCCAACCC CCGGGGTTGG	CACTGACTGC GTGACTGACG	TACTTGTAAG ATGAACATTT	GTAAGTGTTC CATTGACAAG	1140
ATATACTTGA TATATGAACT	TGTAAATGAT ACATTTACTA	AATACCCCAG TTATGGGGTC	CCATCACTAT GGTAGTGATA	TACCCCTCTG ATGGGGAGAC	ACTACTGTAA TGATGACATT	1200
ATGCAGGAGT TACGTCCTCA	TGCCTATATT ACGGATATAA	CCAGAAACAG GGTCTTTGTC	CCACAAAGGA GGTGTTCCTT	GAACCTTTATA CTTGAAATAT	GCTCTGATCA CGAGACTAGT	1260
GCACTACTGA CGTGATGACT	CAGAGCCTCT GTCTCGGAGA	GGATCTAATG CCTAGATTAC	GACAAGTTCTG CTGTTCAAGC	CTGTACTCTT GACATGAGAA	TATGGACATG ATACCTGTAC	1320
AGCACTTTAA TCGTGAAATT	ACTACAGCAA TGATGTCGTT	GCTTATGAGG CGAATACTCC	ACAGTTACAT TGTCATGTGA	GATAGTTACC CTATCAATGG	ACCTCTACTT TGGAGATGAA	1380
TAGACAGGGA ATCTGTCCCT	AAACATAGCA TTTGATTCGT	GCGTACTCTT CGCATGAGAA	TGACAGTAGT ACTGTCATCA	TGCAGAAGAC ACGTCTTCTG	CTTGGCTTCC GAACCGAAGG	1440
CCTCATTGAA GGAGTAACTT	GACCAAAAAG CTGGTTTTTC	TACTACACAG ATGATGTGTC	TCAAGGTTAG AGTTCCAATC	TGATGAGAAT ACTACTCTTA	GACAATGCAC CTGTTACGTG	1500
CTGTATTTTC GACATAAAAG	TAAACCCCAG ATTTGGGGTC	TATGAAGCTT ATACTTCGAA	CTATTCTGGA GATAAGACCT	AAATAATGCT TTTATTACGA	CCAGGCTCTT GGTCCGAGAA	1560
ATATAACTAC TATATTGATG	AGTGATAGCC TCACTATCGG	AGAGACTCTG TCTCTGAGAC	ATAGTGATCA TATCACTAGT	AAATGGCAAA TTTACCGTTT	GTAAATTACA CATTTAATGT	1620
GACTTGTGGA CTGAACACCT	TGCAAAAGTG ACGTTTTTAC	ATGGGCCAGT TACCCGGTCA	CACTAACAAC GTGATTGTTG	ATTTGTTTCT TAAACAAAGA	CTTGATGCGG GAAGTACGCC	1680
ACTCTGGAGT TGAGACCTCA	ATTGAGAGCT TAACTCTCGA	GTTAGGTCTT CAATCCAGAA	TAGACTATGA ATCTGATACT	AAAACTTAAA TTTTGAATTT	CAACTGGATT GTTGACCTAA	1740
TTGAAATTGA AAGTTTAACT	AGCTGCAGAC TCGACGTCTG	AATGGGATCC TTACCCTAGG	CTCAACTCTC GAGTTGAGAG	CACTCGCGTT GTGAGCGCAA	CAACTAAATC GTTGATTTAG	1800
TCAGAAATAGT AGTCTTATCA	TGATCAAAAT ACTAGTTTTA	GATAATTGCC CTATTAACGG	CTGTGATAAC GACACTATTG	TAATCCTCTT ATTAGGAGAA	CTTAATAATG GAATTATTAC	1860
GCTCGGGTGA CGAGCCCACT	AGTTCTGCTT TCAAGACGAA	CCCATCAGCG GGGTAGTCGC	CTCCTCAAAA GAGGAGTTTT	CTATTTAGTT GATAAATCAA	TTCCAGCTCA AAGGTCGAGT	1920
AAGCCGAGGA TTCGGCTCCT	TTCAGATGAA AAGTCTACTT	GGGCACAACT CCCCTGTTGA	CCCAGCTGTT GGGTCGACAA	CTATACCATA GATATGGTAT	CTGAGAGATC GACTCTCTAG	1980
CAAGCAGATT GTTCTGTCTAA	GTTTGCCATT CAAACGGTAA	AACAAAGAAA TTGTTTCTTT	GTGGTGAAAGT CACCACTTCA	GTTCTGTAAA CAAGGACTTT	AAACAATTAA TTTGTTAATT	2040
ACTCTGACCA TGAGACTGGT	TTCAGAGGAC AAGTCTCCTG	TTGAGCATAG AAGTCTGATC	TAGTTGCAGT ATCAACGTCA	GTATGACTTG CATACTGAAC	GGAAGACCTT CCTTCTGGAA	2100
CATTATCCAC GTAATAGGTG	CAATGCTACA GTTACGATGT	GTTAAATTCA CAATTTAAGT	TCCTCACCAG AGGAGTGGCT	CTCTTTTCTT GAGAAAAAGGA	TCTAACGTTG AGATTGCAAC	2160

Fig. 6. (Continuation page 2, SEQ ID NO:6).

AAGTCGTTAT	TTTGCAACCA	TCTGCAGAAG	AGCAGCACCA	GATCGATATG	TCCATTATAT	2220
TTCAGCAATA	AAACGTTGGT	AGACGTCTTC	TCGTCGTGGT	CTAGCTATAC	AGGTAATATA	
TCATTGCAGT	GCTGGCTGGT	GGTTGTGCTT	TGCTACTTTT	GGCCATCTTT	TTTGTGGCCT	2280
AGTAACGTCA	CGACCGACCA	CCAACACGAA	ACGATGAAAA	CCGGTAGAAA	AAACACCGGA	
GTACTTGTA	AAAGAAAGCT	GGTGAATTTA	AGCAGGTACC	TGAACAACAC	GGAACATGCA	2340
CATGAACATT	TTTCTTTTCG	CCACTTAAAT	TCGTCCATGG	ACTTGTGTGT	CCTGTACGT	
ATGAAGAACG	CCTGTTAAGC	ACCCCATCTC	CCCAGTCGGT	CTCTTCTTCT	TTGTCTCAGT	2400
TACTTCTTGC	GGACAATTCT	TGGGGTAGAG	GGGTCAGCCA	GAGAAGAAGA	AACAGAGTCA	
CTGAGTCATG	CCAACCTCTC	ATCAATACTG	AATCTGAGAA	TTGCAGCGTG	TCCTCTAACC	2460
GACTCAGTAC	GGTTGAGAGG	TAGTTATGAC	TTAGACTCTT	AACGTCGCAC	AGGAGATTGG	
AAGAGCAGCA	TCAGCAAACA	GGCATAAAGC	ACTCCATCTC	TGTACCATCT	TATCACACAT	2520
TTCTCGTCGT	AGTCGTTTGT	CCGTATTTCG	TGAGGTAGAG	ACATGGTAGA	ATAGTGTGTA	
CTGGTTGGCA	CCTGGACAAT	TGTGCAATGA	GCATAAGTGG	ACATTCTCAC	ATGGGGCACA	2580
GACCAACCGT	GGACCTGTTA	ACACGTTACT	CGTATTACCC	TGTAAGAGTG	TACCCCGTGT	
TTAGTACAAA	GGTACAGTGG	GCAAAGGAGA	TAGTGACTTC	AATGACAGTG	ACTCTGATAC	2640
AATCATGTTT	CCATGTCACC	CGTTTCCTCT	ATCACTGAAG	TTACTGTAC	TGAGACTATG	
TAGTGGAGAA	TCAGAAAAGA	AGAGCATTGA	GCAGCCAATG	CAGGCACAAG	CCAGTGCTCA	2700
ATCACCTCTT	AGTCTTTTCT	TCTCGTAACT	CGTCGGTTAC	GTCCGTGTTC	GGTCACGAGT	
ATACACAGAT	GAATCAGCAG	GGTTCCGACA	TGCCGATAAC	TATTTTCAGCC	ACCGAATCAA	2760
TATGTGTCTA	CTTAGTCGTC	CCAAGGCTGT	ACGGCTATTG	ATAAAGTCGG	TGGCTTAGTT	
CAAGGGTCCA	GAAAATGGGA	ACTGCACATT	GCAATATGAA	AAGGGCTATA	GA CTGTCTTA	2820
GTTCCAGGT	CTTTTACCTT	TGACGTGTAA	CGTTATAC TT	TTCCCGATAT	CTGACAGAA T	
CTCTGTAGCT	CCTGTATATT	ACAATACCTA	CCATGCAAGA	ATGCCTAACC	TGCACATACC	2880
GAGACATCGA	GGACATATAA	TGTTATGGAT	GGTACGTTCT	TACGGATTGG	ACGTGTATGG	
GAACCATACC	CTTAGAGACC	CTTATTACCA	TATCAATAAT	CCTGTTGCTA	ATCGGATGCA	2940
CTTGGTATGG	GAATCTCTGG	GAATAATGGT	ATAGTTATTA	GGACAAOGAT	TAGCCTACGT	
GGCGGAATAT	GAAAGAGATT	TAGTCAACAG	AAGTGCAACG	TTATCTCCGC	AGAGATCGTC	3000
CCGCCTTATA	CTTTCTCTAA	ATCAGTTGTC	TTCACGTTGC	AATAGAGGCG	TCTCTAGCAG	
TAGCAGATAC	CAAGAATTCA	ATTACAGTCC	GCAGATATCA	AGACAGCTTC	ATCCTTCAGA	3060
ATCGTCTATG	GTTCTTAAGT	TAATGTCAGG	CGTCTATAGT	TCTGTGCAAG	TAGGAAGTCT	
AATTGCTACA	ACCTTTTAAT	CATTAGGCAT	GCAAGTGAGA	ATGCACAAAG	GCAAGTGCTT	3120
TTAACGATGT	TGGAAATTA	GTAATCCGTA	CGTTCACTCT	TACGTGTTTC	CGTTCACGAA	
TAGCATGAAA	GCTAAATATA	TGGAGTCTCC	CCTTTCCCTC	TGATGGATGG	GGGGAGACAC	3180
ATCGTACTTT	CGATTTATAT	ACCTCAGAGG	GGAAAGGGAG	ACTACCTACC	CCCCTCTGTG	
AGGACAGTGC	ATAAATATAC	AGCTGCTTTC	TATTTGCATT	TCAC TTGGGA	ATTTTTTGTT	3240
TCCTGTCACG	TATTTATATG	TCGACGAAAG	ATAAACGTAA	AGTGAAC CCT	TAAAAAACAA	
TTTTTTACAT	ATTTATTTTT	CCTGAATTGA	ATGTGACATT	GTCCTGTCAC	CTAACTAGCA	3300
AAAAAATGTA	TAAATAAAAA	GGACTTAACT	TACACTGTAA	CAGGACAGTG	GATTGATCGT	

Fig. 6. (Continuation page 3, SEQ ID NO:6).

MVCCGPGRML LGWAGLLVLA ALCLLQVPGA QAAACEPVRI PLCKSLPWNM TKMPNHLHHS 60
TQANAILAME QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPCKS VCERARQGCE 120
PILIKYRHSW PESLACDELP VYDRGVCISP EAIVTADGAD FPMDSSTGHC RGASSERCKC 180
KPV RATQKTY FRNNYNYVIR AKVKEVKMKC HDVTAVVEVK EILKASLVNI PRDTVNLYTT 240
SGCLCPPLTV NEEYVIMGYE DEERSRLLLV EGSIAEKWKD RLGKKVKRWD MKLRHLGLGK 300
TDASDSTQNQ KSGRNSNPRP ARS.

Figure 7. Deduced amino acid sequence of mouse FRZB-1 protein. SEQ ID NO:7.

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 8. Nucleotide sequence of the full-length mouse FRZB-1 cDNA. SEQ ID NO:8.

AAGCCTGGGA	CCATGGTCTG	CTGCGGCCCG	GGACGGATGC	TGCTAGGATG	GGCCGGGTTG	60
TTCGGACCCT	GGTACCAGAC	GACGCCGGGC	CCTGCCTACG	ACGATCCTAC	CCGGCCCAAC	
CTAGTCCTGG	CTGCTCTCTG	CCTGCTCCAG	GTGCCCCGAG	CTCAGGCTGC	AGCCTGTGAG	120
GATCAGGACC	GACGAGAGAC	GGACGAGGTC	CACGGGCCTC	GAGTCCGACG	TCGGACACTC	
CCTGTCCGCA	TCCCCGCTGTG	CAAGTCCCTT	CCCTGGAACA	TGACCAAGAT	GCCCAACCAC	180
GGACAGGCGT	AGGGCGACAC	GTTCAGGGAA	GGGACCTTGT	ACTGGTTCTA	CGGGTTGGTG	
CTGCACCACA	GCACCCAGGC	TAACGCCATC	CTGGCCATGG	AACAGTTCGA	AGGGCTGCTG	240
GACGTGGTGT	CGTGGGTCCG	ATTGCGGTAG	GACCGGTACC	TTGTCAAGCT	TCCCCACGAC	
GGCACCCACT	GCAGCCCGGA	TCTTCTCTTC	TTCCTCTGTG	CAATGTACGC	ACCCATTTGC	300
CCGTGGGTGA	CGTCGGGCCT	AGAAGAGAAG	AAGGAGACAC	GTTACATGCG	TGGGTAAACG	
ACCATCGACT	TCCAGCACGA	GCCCATCAAG	CCCTGCAAGT	CTGTGTGTGA	GCGCGCCCCGA	360
TGGTAGCTGA	AGGTCGTGCT	CGGGTAGTTC	GGGACGTTCA	GACACACACT	CGCGCGGGCT	
CAGGGCTGCG	AGCCCATTTCT	CATCAAGTAC	CGCCACTCGT	GGCCGGAAG	CTTGGCCTGC	420
GTCCCAGACG	TCGGGTAAGA	GTAGTTCATG	GCGGTGAGCA	CCGGCCTTTC	GAACCGGACG	
GACGAGCTGC	CGGTGTACGA	CCGCGGCGTG	TGCATCTCTC	CTGAGGCCAT	CGTCACCGCG	480
CTGCTCGACG	GCCACATGCT	GGCGCCGCAC	ACGTAGAGAG	GACTCCGGTA	GCAGTGGCGC	
GACGGAGCGG	ATTTTCCTAT	GGATTCAAGT	ACTGGACACT	GCAGAGGGGC	AAGCAGCGAA	540
CTGCCTCGCC	TAAAAGGATA	CCTAAGTTCA	TGACCTGTGA	CGTCTCCCCG	TTCGTCGCTT	
CGTTGCAAAT	GTAAGCCTGT	CAGAGCTACA	CAGAAGACCT	ATTTCCGGAA	CAATTACAAC	600
GCAACGTTTA	CATTTCGACA	GTCTCGATGT	GTCTTCTGGA	TAAAGGCCTT	GTTAATGTTG	
TATGTCATCC	GGGCTAAAGT	TAAAGAGGTA	AAGATGAAAT	GTCATGATGT	GACCGCCGTT	660
ATACAGTAGG	CCCGATTTCA	ATTTCTCCAT	TTCTACTTTA	CAGTACTACA	CTGGCGGCAA	
GTGGAAGTGA	AGGAAATTCT	AAAGGCATCA	CTGGTAAACA	TTCCAAGGGA	CACCGTCAAT	720
CACCTTCACT	TCCTTTAAGA	TTTCCGTAGT	GACCATTTGT	AAGGTTCCCT	GTGGCAGTTA	
CTTTATACCA	CCTCTGGCTG	CCTCTGTCCT	CCACTTACTG	TCAATGAGGA	ATATGTCATC	780
GAAATATGGT	GGAGACCGAC	GGAGACAGGA	GGTGAATGAC	AGTTACTCCT	TATACAGTAG	
ATGGGCTATG	AAGACGAGGA	ACGTTCCAGG	TTACTCTTGG	TAGAAGGCTC	TATAGCTGAG	840
TACCCGATAC	TTCTGCTCCT	TGCAAGGTCC	AATGAGAACC	ATCTTCCGAG	ATATCGACTC	
AAGTGGAAGG	ATCGGCTTGG	TAAGAAAGTC	AAGCGCTGGG	ATATGAAACT	CCGACACCTT	900
TTCACCTTCC	TAGCCGAACC	ATTCTTTTCA	TTCGCGACCC	TATACTTTGA	GGCTGTGGAA	
GGACTGGGTA	AAACTGATGC	TAGCGATTCC	ACTCAGAATC	AGAAGTCTGG	CAGGAACCTC	960
CCTGACCCAT	TTTGACTACG	ATCGCTAAGG	TGAGTCTTAG	TCTTCAGACC	GTCCTTGAGA	

<p> AATCCCCGGC TTAGGGGCCG </p>	<p> CAGCACGCAG GTCGTGCGTC </p>	<p> CTAAATCCTG GATTTAGGAC </p>	<p> AAATGTAAAA TTTACATTTT </p>	<p> GGCCACACCC CCGGTGTGGG </p>	<p> ACGGACTCCC TGCCTGAGGG </p>	1020
<p> TTCTAAGACT AAGATTCTGA </p>	<p> GGCGCTGGTG CCGCGACCAC </p>	<p> GACTAACAAA CTGATTGTTT </p>	<p> GGAAAACCGC CCTTTTGGCG </p>	<p> ACAGTTGTGC TGTCAACACG </p>	<p> TCGTGACCGA AGCACTGGCT </p>	1080
<p> TTGTTTACCG AACAAATGGC </p>	<p> CAGACACCGC GTCTGTGGCG </p>	<p> GTGGCTACCG CACCGATGGC </p>	<p> AAGTTACTTC TTCAATGAAG </p>	<p> CGGTCCCCTT GCCAGGGGAA </p>	<p> TCTCCTGCTT AGAGGACGAA </p>	1140
<p> CTTAATGGCG GAATTACCGC </p>	<p> TGGGGTTAGA ACCCCAATCT </p>	<p> TCCTTTAATA AGGAAATTAT </p>	<p> TGTTATATAT ACAATATATA </p>	<p> TCTGTTTCAT AGACAAAGTA </p>	<p> CAATCACGTG GTTAGTGCAC </p>	1200
<p> GGGACTGTTT CCCTGACAAG </p>	<p> TTTTGCAACC AAAACGTTGG </p>	<p> AGAATAGTAA TCTTATCATT </p>	<p> ATTAAATATG TAATTTATAC </p>	<p> TTGATGCTAA AACTACGATT </p>	<p> GGT'TTCTGTA CCAAAGACAT </p>	1260
<p> CTGGACTCCC GACCTGAGGG </p>	<p> TGGGTTTAAT ACCCAAATTA </p>	<p> TTGGTGTTCT AACCACAAGA </p>	<p> GTACCCTGAT CATGGGACTA </p>	<p> TGAGAATGCA ACTCTTACGT </p>	<p> ATGTTTTCATG TACAAAGTAC </p>	1320
<p> TAAAGAGAGA ATTTCTCTCT </p>	<p> ATCCTGGTCA TAGGACCAGT </p>	<p> TATCTCAAGA ATAGAGTTCT </p>	<p> ACTAGATATT TGATCTATAA </p>	<p> GCTGTAAGAC CGACATTCTG </p>	<p> AGCCTCTGCT TCGGAGACGA </p>	1380
<p> GCTGCGCTTA CGACGCGAAT </p>	<p> TAGTCTTGTG ATCAGAACAC </p>	<p> TTTGTATGCC AAACATACGG </p>	<p> TTTGTCCATT AAACAGGTAA </p>	<p> TCCCTCATGC AGGGAGTACG </p>	<p> TGTGAAAGTT ACACTT'TCAA </p>	1440
<p> ATACATGTTT TATGTACAAA </p>	<p> ATAAAGGTAG TATTTCCATC </p>	<p> AACGGCATTT TTGCCGTAAA </p>	<p> TGAAATCAGA ACTTTAGTCT </p>	<p> CACTGCACAA GTGACGTGTT </p>	<p> GCAGAGTAGC CGTCTCATCG </p>	1500
<p> CCAACACCAG GGTTGTGGTC </p>	<p> GAAGCATTTA CTTCGTAAAT </p>	<p> TGAGGAAACG ACTCCTTTGC </p>	<p> CCACACAGCA GGTGTGTCGT </p>	<p> TGACTTATTT ACTGAATAAA </p>	<p> TCAAGATTGG AGTTCTAACC </p>	1560
<p> CAGGCAGCAA GTCCGTCGTT </p>	<p> AATAAATAGT TTATTTATCA </p>	<p> GTTGGGAGCC CAACCCTCGG </p>	<p> AAGAAAAGAA TTCTTTTCTT </p>	<p> TATTTTGCCT ATAAAACGGA </p>	<p> GGTTAAGGGG CCAATTCCCC </p>	1620
<p> CACACTGGAA GTGTGACCTT </p>	<p> TCAGTAGCCC AGTCATCGGG </p>	<p> TTGAGCCATT AACTCGGTAA </p>	<p> AACAGCAGTG TTGTGCTCAC </p>	<p> TTCTTCTGGC AAGAAGACCG </p>	<p> AAGTTT'TTGA TTCAAAACT </p>	1680
<p> TTTGTTTATA AAACAAGTAT </p>	<p> AATGTATTCA TTACATAAGT </p>	<p> CGAGCATTAG GCTCGTAATC </p>	<p> AGATGAACTT TCTACTTGAA </p>	<p> ATAACTAGAC TATTGATCTG </p>	<p> ATCTGTTGTT TAGACAACAA </p>	1740
<p> ATCTCTATAG TAGAGATATC </p>	<p> CTCTGCTTCC GAGACGAAGG </p>	<p> TTCTAAAATCA AAGATTTAGT </p>	<p> AACCCATTGT TTGGGTAACA </p>	<p> TGGATGCTCC ACCTACGAGG </p>	<p> CTCTCCATTTC GAGAGGTAAG </p>	1800

MVCGSPGGML	LLRAGLLALA	ALCLLRVPGA	RAACEPVRI	PLCKSLPWNM	TKMPNHLHHS	60
TQANAILAIE	QFEGLLGTHC	SPDLLFFLCA	MYAPICTIDF	QHEPIKPCKS	VCERARQGCE	120
PILIKYRHSW	PENLACEELP	VYDRGVCISP	EAIVTADGAD	FPMDSNGNC	RGASSERCKC	180
KPIRATQKTY	FRNNYNYVIR	AKVKEIKTKC	HDVTAVVEVK	EILKSSLVNI	PRDTVNLYTS	240
SGCLCPPLNV	NEEYIIMGYE	DEERSRLLL	EGSIAEKWKD	RLGKKVKRWD	MKLRLHLGLSK	300
SDSSNSDSTQ	SQKSGRNSNP	ROARN.				

Figure 9. Deduced amino acid sequence of human FRZB-1 protein. SEQ ID NO:9.

Figure 10. Nucleotide sequence of the full-length human FRZB-1 cDNA. SEQ ID NO:10.
This sequence was assembled from public ESTs from the Genbank database
(accession numbers: H18848, R63748, W38677, W44760, H38379 and N71244).

GGCGGAGCGG	GCCTTTTGGC	GTCCACTGCG	CGGCTGCACC	CTGCCCCATC	TGCCGGGATC	60
CCGCCTCGCC	CGGAAAACCG	CAGGTGACGC	GCCGACGTGG	GACGGGGTAG	ACGGCCCTAG	
ATGGTCTGCG	GCAGCCCGGG	AGGGATGCTG	CTGCTGCGGG	CCGGGCTGCT	TGCCCTGGCT	120
TACCAGACGC	CGTCGGGCCC	TCCCTACGAC	GACGACGCCC	GGCCCCGACG	ACGGGACCGA	
GCTCTCTGCC	TGCTCCGGGT	GCCCCGGGGCT	CGGGCTGCAG	CCTGTGAGCC	CGTCCGCATC	180
CGAGAGACGG	ACGAGGCCCA	CGGGCCCCGA	GCCCCGACGTC	GGACACTCGG	GCAGGCGTAG	
CCCCTGTGCA	AGTCCCTGCC	CTGGAACATG	ACTAAGATGC	CCAACCACCT	GCACCACAGC	240
GGGGACACGT	TCAGGGACGG	GACCTTGTA	TGATTCTACG	GGTTGGTGGA	CGTGGTGTCTG	
ACTCAGGCCA	ACGCCATCCT	GGCCATCGAG	CAGTTCGAAG	GTCTGCTGGG	CACCCACTGC	300
TGAGTCCGGT	TGCGGTAGGA	CCGGTAGCTC	GTCAAGCTTC	CAGACGACCC	GTGGGTGACG	
AGCCCCGATC	TGCTCTTCTT	CCTCTGTGCC	ATGTACGCGC	CCATCTGCAC	CATTGACTTC	360
TCGGGGCTAG	ACGAGAAGAA	GGAGACACGG	TACATGCGCG	GGTAGACGTG	GTAAGTGAAG	
CAGCACGAGC	CCATCAAGCC	CTGTAAGTCT	GTGTGCGAGC	GGGCCCCGCA	GGGCTGTGAG	420
GTCGTGCTCG	GGTAGTTCGG	GACATTGACA	CACACGCTCG	CCCGGGCCGT	CCCGACACTC	
CCCATACTCA	TCAAGTACCG	CCACTCGTGG	CCGGAGAACC	TGGCCTGCGA	GGAGCTGCCA	480
GGGTATGAGT	AGTTCATGGC	GGTGAGCACC	GGCCTCTTGG	ACCGGACGCT	CCTCGACGGT	
GTGTACGACA	GGGGCGTGTG	CATCTCTCCC	GAGGCCATCG	TACTGCGGGA	CGGAGCTGAT	540
CACATGCTGT	CCCCGCACAC	GTAGAGAGGG	CTCCGGTAGC	AATGACGCCT	GCCTCGACTA	
TTTCCTATGG	ATTCTAGTAA	CGGAAACTGT	AGAGGGGCAA	GCAGTGAACG	CTGTAAATGT	600
AAAGGATACC	TAAGATCATT	GCCTTTGACA	TCTCCCCGTT	CGTCACTTGC	GACATTTACA	
AAGCCTATTA	GAGCTACACA	GAAGACCTAT	TTCCGGAACA	ATTACAACCTA	TGTCATTTCGG	660
TTCGGATAAT	CTCGATGTGT	CTTCTGGATA	AAGGCCTTGT	TAATGTTGAT	ACAGTAAGCC	
GCTAAAGTTA	AAGAGATAAA	GACTAAGTGC	CATGATGTGA	CTGCAGTAGT	GGAGGTGAAG	720
CGATTTCAAT	TTCTCTATTT	CTGATTACAG	GTACTACACT	GACGTCATCA	CCTCCACTTC	
GAGATTCTAA	AGTCCTCTCT	GGTAAACATT	CCACGGGACA	CTGTCAACCT	CTATACCAGC	780
CTCTAAGATT	TCAGGAGAGA	CCATTTGTAA	GGTGCCCTGT	GACAGTTGGA	GATATGGTCTG	
TCTGGCTGCC	TCTGCCCTCC	ACTTAATGTT	AATGAGGAAT	ATATCATCAT	GGGCTATGAA	840
AGACCGACGG	AGACGGGAGG	TGAATTACAA	TTACTCCTTA	TATAGTAGTA	CCCGATACTT	

GATGAGGAAC	GTTCCAGATT	ACTCTTGGTG	GAAGGCTCTA	TAGCTGAGAA	GTGGAAGGAT	900
CTACTCCTTG	CAAGGTCTAA	TGAGAACCAC	CTTCCGAGAT	ATCGACTCTT	CACCTTCCTA	
CGACTCGGTA	AAAAAGTTAA	GCGCTGGGAT	ATGAAGCTTC	GTCATCTTGG	ACTCAGTAAA	960
GCTGAGCCAT	TTTTTCAATT	CGCGACCCTA	TACTTCGAAG	CAGTAGAACC	TGAGTCATTT	
AGTGATTCTA	GCAATAGTGA	TTCCACTCAG	AGTCAGAAAGT	CTGGCAGGAA	CTCGAACCCC	1020
TCACTAAGAT	CGTTATCACT	AAGGTGAGTC	TCAGTCTTCA	GACCGTCCTT	GAGCTTGGGG	
CGGCAAGCAC	GCAACTAAAT	CCCGAAATAC	AAAAAGTAAC	ACAGTGGACT	TCCTATTAAG	1080
GCCGTTTCGTG	CGTTGATTTA	GGGCTTTATG	TTTTTCATTG	TGTCACCTGA	AGGATAATTC	
ACTTACTTGC	ATTGCTGGAC	TAGCAAAGGA	AAATTGCACT	ATTGCACATC	ATATTCTATT	1140
TGAATGAACG	TAACGACCTG	ATCGTTTCTT	TTTAACGTGA	TAACGTGTAG	TATAAGATAA	
GTTTACTATA	AAAATCATGT	GATAACTGAT	TATTACTTCT	GTTTCTCTTT	TGGTTTCTGC	1200
CAAATGATAT	TTTTAGTACA	CTATTGACTA	ATAATGAAGA	CAAAGAGAAA	ACCAAAGACG	
TTCTCTCTTC	TCTCAACCCC	TTTGTAATGG	TTTGGGGGCA	GA CTCTTAAG	TATATTGTGA	1260
AAGAGAGAAG	AGAGTTGGGG	AAACATTACC	AAACCCCGCT	CTGAGAATTC	ATATAACACT	
GTTTTCTATT	TCACTAATCA	TGAGAAAAAC	TGTTCTTTTG	CAATAATAAT	AAATTAAACA	1320
CAAAAGATAA	AGTGATTAGT	ACTCTTTTTG	ACAAGAAAAC	GTTATTATTA	TTTAATTTGT	
TGCTGTTACC	AGAGCCTCTT	TGCTGAGTCT	CCAGATGTTA	ATTTACTTTC	TGCACCCCAA	1380
ACGACAATGG	TCTCGGAGAA	ACGACTCAGA	GGTCTACAAT	TAAATGAAAG	ACGTGGGGTT	
TTGGGAATGC	AATATTGGAT	GAAAAGAGAG	GTTTCTGGTA	TTACAGAAA	GCTAGATATG	1440
AACCTTACG	TTATAACCTA	CTTTTCTCTC	CAAAGACCAT	AAGTGTCTTT	CGATCTATAC	
CCTTAAACA	TACTCTGCCG	ATCTAATTAC	AGCCTTATTT	TTGTATGCCT	TTTGGGCATT	1500
GGAATTTTGT	ATGAGACGGC	TAGATTAATG	TCGGAATAAA	AACATACGGA	AAACCCGTAA	
CTCCTCATGC	TTAGAAAGTT	CCAAATGTTT	ATAAAGGTAA	AATGGCAGTT	TGAAGTCAAA	1560
GAGGAGTACG	AATCTTTCAA	GGTTTACAAA	TATTTCCATT	TTACCGTCAA	ACTTCAGTTT	
TGTCACATAG	GCAAAGCAAT	CAAGCACCAG	GAAGTGTTTA	TGAGGAAACA	ACACCCAAGA	1620
ACAGTGATATC	CGTTTCGTTA	GTTTCGTGGTC	CTTCACAAAT	ACTCCTTTGT	TGTGGGTCTT	
TGAATTATTT	TTGAGACTGT	CAGGAAGTAA	AATAAATAGG	AGCTTAAGAA	AGAACATTTT	1680
ACTTAATAAA	AACCTTGACA	GTCCTTCATT	TTATTTATCC	TCGAATTCTT	TCTTGTAATAA	
GCCTGATTGA	GAAGCACAAC	TGAAACCAGT	AGCCGCTGGG	GTGTTAATGG	TAGCATTCTT	1740
CGGACTAACT	CTTCGTGTTG	ACTTTGGTCA	TCGGCGACCC	CACAATTACC	ATCGTAAGAA	
CTTTTGGCAA	TACATTTGAT	TTGTTTCATGA	ATATATTAAT	CAGCATTAGA	GAAATGAATT	1800
GAAAACCGTT	ATGTAAACTA	AACAAGTACT	TATATAATTA	GTCGTAATCT	CTTTACTTAA	
ATAACTAGAC	ATCTGCTGTT	ATCACCATAG	TTTTGTTTAA	TTTGCTTCCT	TTTAAATAAA	1860
TATTGATCTG	TAGACGACAA	TAGTGGTATC	AAAACAAATT	AAACGAAGGA	AAATTTATTT	
CCCATTGGTG	AAAGTCAAAA	AAAAAAAAAA	AAA			
GGGTAACCAC	TTTCAGTTTT	TTTTTTTTTT	TTT			